solution

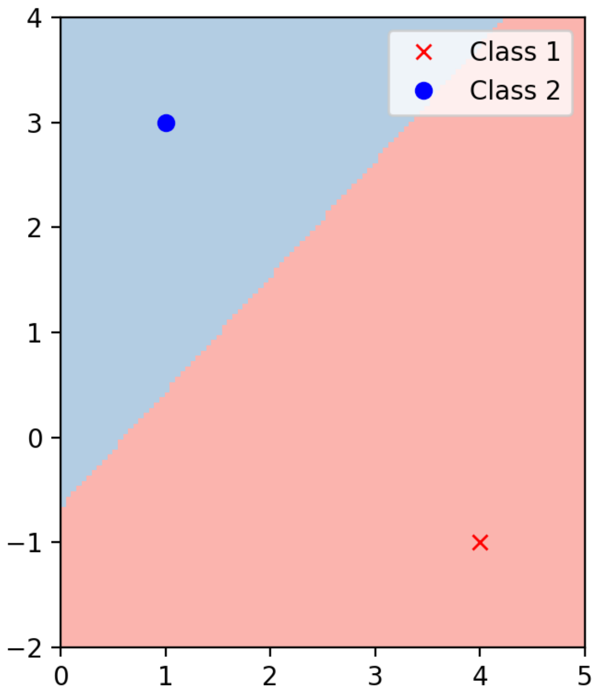
Question 1

(a)

We can derive the decision rule following the Bayes rule ($\underline{X} = [x_1, x_2]^T$):

$$\begin{aligned}
P(S_i|X=x) &\propto P(X=x|S_i)P(S_i), i=1,2 \\
\ln\frac{P(S_1|X=x)}{P(S_2|X=x)} &= \ln\frac{P(X=x|S_1)P(S_1)}{P(X=x|S_2)P(S_2)} \\
&= \ln\frac{N(\underline{x}, \underline{m_1}, \underline{\underline{\Sigma_1}})}{N(\underline{x}, \underline{m_2}, \underline{\underline{\Sigma_2}})} \\
&= \ln\frac{|\underline{\underline{\Sigma_1}}|^{-1/2}\exp\left(-\frac{1}{2}(\underline{x}-\underline{m_1})^T\underline{\underline{\Sigma_1}}^{-1}(\underline{x}-\underline{m_1})\right)}{|\underline{\underline{\Sigma_2}}|^{-1/2}\exp\left(-\frac{1}{2}(\underline{x}-\underline{m_2})^T\underline{\underline{\Sigma_2}}^{-1}(\underline{x}-\underline{m_2})\right)} \\
&= -\frac{1}{2}[(\underline{x}-\begin{bmatrix}1\\-1\end{bmatrix})^T\begin{bmatrix}2 & 3\\3 & 6.5\end{bmatrix}^{-1}(\underline{x}-\begin{bmatrix}1\\-1\end{bmatrix})-(\underline{x}-\begin{bmatrix}2\\3\end{bmatrix})^T\begin{bmatrix}2 & 3\\3 & 6.5\end{bmatrix}^{-1}(\underline{x}-\begin{bmatrix}2\\3\end{bmatrix})] \\
&= 1.375x_1 - 1.25x_2 - 0.8125
\end{aligned}$$

So we have decision rule that if $\ln\frac{P(S_1|X=x)}{P(S_2|X=x)} > 0$, then x belongs to class 1, else x belongs to class 2.

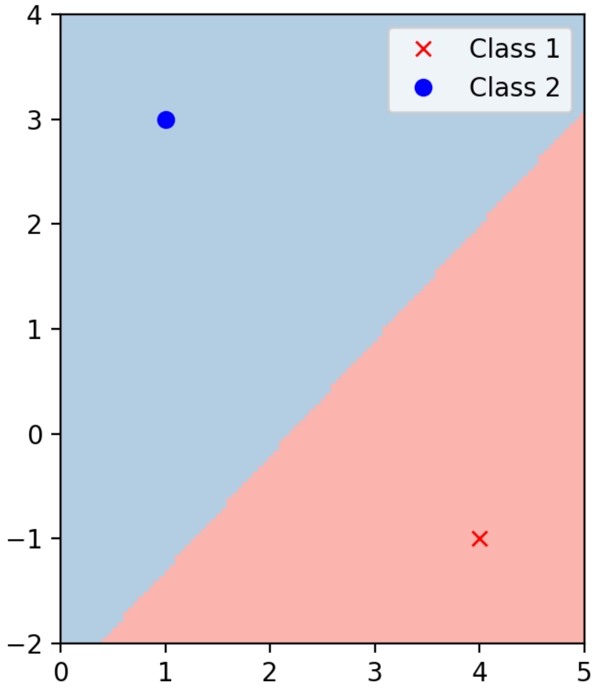


(b)

We can derive the decision rule following the Bayes rule:

$$
\begin{aligned}
P(S_i|X=x) &\propto P(X=x|S_i)P(S_i), i=1,2 \\
\ln\frac{P(S_1|X=x)}{P(S_2|X=x)} &= \ln\frac{P(X=x|S_1)P(S_1)}{P(X=x|S_2)P(S_2)} \\
&= \ln\frac{N(\underline{x},\underline{m_1},\underline{\underline{\Sigma_1}})P(S_1)}{N(\underline{x},\underline{m_2},\underline{\underline{\Sigma_2}})P(S_2)} \\
&= \ln\frac{P(S_1)|\underline{\underline{\Sigma_1}}|^{-1/2}\exp\left(-\frac{1}{2}(\underline{x}-\underline{m_1})^T\underline{\underline{\Sigma_1}}^{-1}(\underline{x}-\underline{m_1})\right)}{P(S_2)|\underline{\underline{\Sigma_2}}|^{-1/2}\exp\left(-\frac{1}{2}(\underline{x}-\underline{m_2})^T\underline{\underline{\Sigma_2}}^{-1}(\underline{x}-\underline{m_2})\right)} \\
&= \ln\frac{1}{9}+\left(-\frac{1}{2}[(\underline{x}-\begin{bmatrix}1\\-1\end{bmatrix})^T\begin{bmatrix}2&3\\3&6.5\end{bmatrix}^{-1}(\underline{x}-\begin{bmatrix}1\\-1\end{bmatrix})-(\underline{x}-\begin{bmatrix}2\\3\end{bmatrix})^T\begin{bmatrix}2&3\\3&6.5\end{bmatrix}^{-1}(\underline{x}-\begin{bmatrix}2\\3\end{bmatrix})]\right) \\
&= 1.375x_1-1.25x_2-3.00972
\end{aligned}
$$

So we have decision rule that if $\ln\frac{P(S_1|X=x)}{P(S_2|X=x)}>0$, then x belongs to class 1, else x belongs to class 2.



Question 2

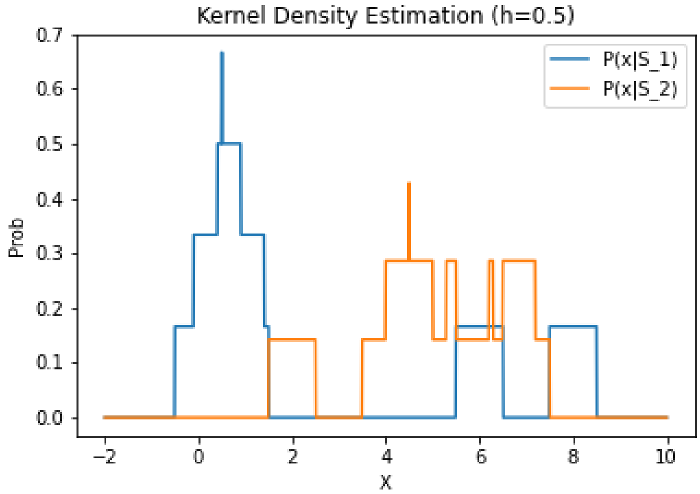
(a)

h=0.5,

$$
\phi(x)=\begin{cases}1, & -0.5\le x\le 0.5\\ 0, & \text{otherwise}\end{cases}
$$

$$
P(x|S_1)=\frac{1}{6}\sum_{i=1}^{6}\phi(x-x_i)
$$

$$
P(x|S_2)=\frac{1}{7}\sum_{i=1}^{7}\phi(x-x_i)
$$



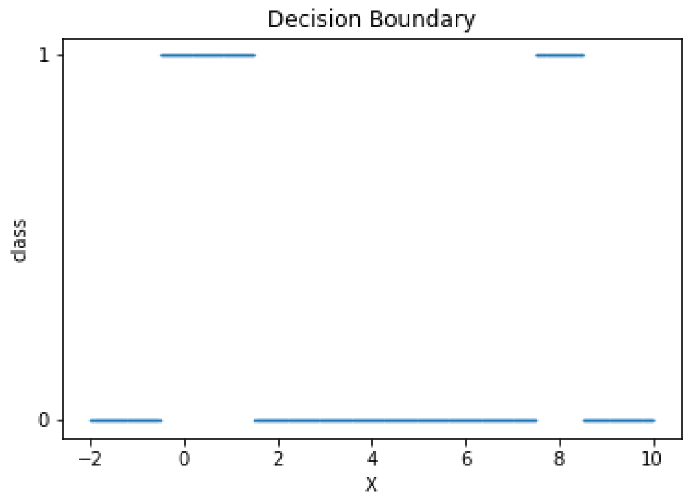
(b)

The prior is $P(S_1) = \frac{6}{13}, p(S_2) = \frac{7}{13}$

(c)

Following Bayes rule, we can derive the decision rule as follows:

$$\begin{aligned}\frac{P(S_1|X=x)}{P(S_2|X=x)} &= \frac{P(X=x|S_1)P(S_1)}{P(X=x|S_2)P(S_2)} \\ &= \frac{\sum_{i=1}^{n_1}\phi(x-x_i)}{\sum_{i=1}^{n_2}\phi(x-x_i)}\end{aligned}$$



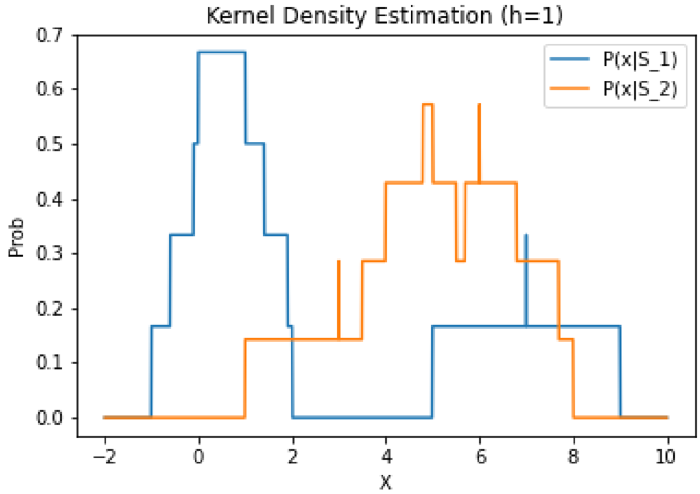
(d)

h=1,

$$\phi(x) = \begin{cases} 1, & -1 \le x \le 1 \\ 0, & \text{otherwise} \end{cases}$$

$$P(x|S_1) = \frac{1}{6}\sum_{i=1}^{6}\phi(x-x_i)$$

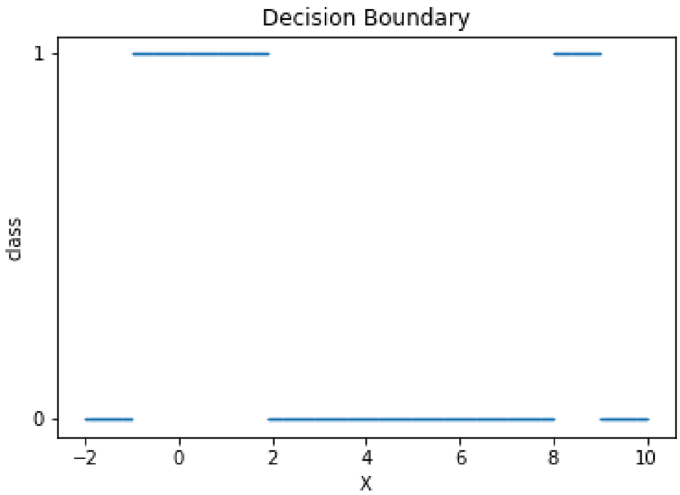
$$P(x|S_2) = \frac{1}{7}\sum_{i=1}^{7}\phi(x-x_i)$$



The prior is $P(S_1) = \frac{6}{13}, p(S_2) = \frac{7}{13}$

Following Bayes rule, we can derive the decision rule as follows:

$$\begin{aligned}\frac{P(S_1|X=x)}{P(S_2|X=x)} &= \frac{P(X=x|S_1)P(S_1)}{P(X=x|S_2)P(S_2)} \\ &= \frac{\sum_{i=1}^{n_1}\phi(x-x_i)}{\sum_{i=1}^{n_2}\phi(x-x_i)}\end{aligned}$$



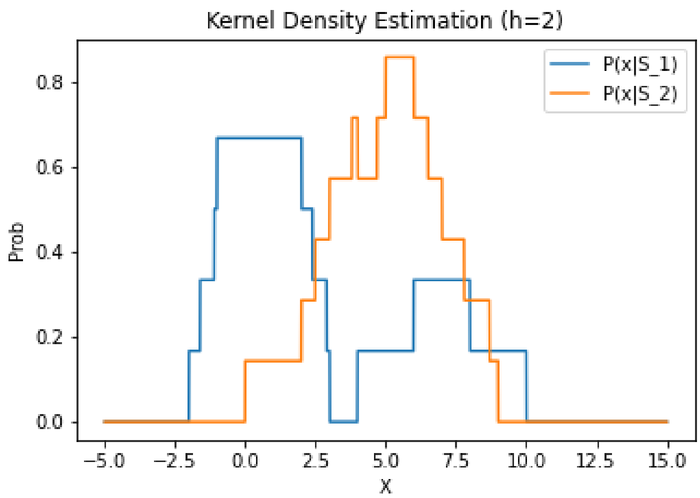
(e)

h=2,

$$\phi(x) = \begin{cases} 1, & -2 \le x \le 2 \\ 0, & \text{otherwise} \end{cases}$$

$$P(x|S_1) = \frac{1}{6}\sum_{i=1}^{6}\phi(x-x_i)$$

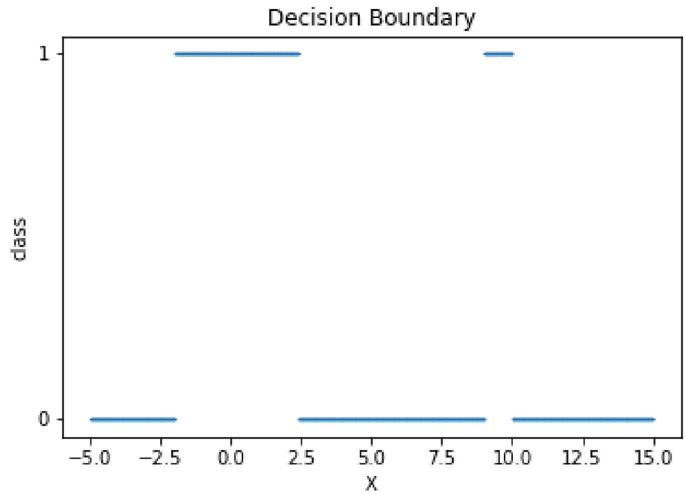
$$P(x|S_2) = \frac{1}{7}\sum_{i=1}^{7}\phi(x-x_i)$$



The prior is $P(S_1) = \frac{6}{13}, p(S_2) = \frac{7}{13}$

Following Bayes rule, we can derive the decision rule as follows:

$$\begin{aligned}\frac{P(S_1|X=x)}{P(S_2|X=x)} &= \frac{P(X=x|S_1)P(S_1)}{P(X=x|S_2)P(S_2)} \\ &= \frac{\sum_{i=1}^{n_1}\phi(x-x_i)}{\sum_{i=1}^{n_2}\phi(x-x_i)}\end{aligned}$$



Question 3

(a)

For $x = 0$, we can compute the distance of each point to it as follows:

$d_1 = 0.9, d_2 = 0.7, d_3 = 0.5, d_4 = 0.3, d_5 = 0.1, d_6 = 0.1, d_7 = 0.3, d_8 = 0.5, d_9 = 0.7, d_{10} = 0.9$

The nearest 4 points are $(-0.1, 0.01), (0.1, 0.01), (-0.3, 0.09), (0.3, 0.09)$

Thus, the estimated $\hat{y}(x) = \frac{0.01+0.01+0.09+0.09}{4} = 0.05$

For $x = 0.4$, we can compute the distance of each point to it as follows:

$d_1 = 1.3, d_2 = 1.1, d_3 = 0.9, d_4 = 0.7, d_5 = 0.5, d_6 = 0.3, d_7 = 0.1, d_8 = 0.1, d_9 = 0.3, d_{10} = 0.5$

The nearest 4 points are $(0.3, 0.09), (0.5, 0.25), (0.1, 0.01), (0.7, 0.49)$

Thus, the estimated $\hat{y}(x) = \frac{0.09+0.25+0.01+0.49}{4} = 0.21$

(b)

$y(0) = 0, y(0.4) = 0.16$

Therefore, the mean square error is $MSE = \frac{1}{2}((0.05-0)^2 + (0.21-0.16)^2) = 0.0025$

(c)

$$k(\underline{x}, \underline{x_n}) = 1 - \frac{d(\underline{x}, \underline{x_n})}{d_{\max}}$$

For $x = 0$, we can apply the kernel regression as follows:

$d_1 = 0.9, d_2 = 0.7, d_3 = 0.5, d_4 = 0.3, d_5 = 0.1, d_6 = 0.1, d_7 = 0.3, d_8 = 0.5, d_9 = 0.7, d_{10} = 0.9$

$d_{\max} = 0.5$

$k_1 = -0.8, k_2 = -0.4, k_3 = 0, k_4 = 0.4, k_5 = 0.8, k_6 = 0.8, k_7 = 0.4, k_8 = 0, k_9 = -0.4, k_{10} = -0.8$

$$\hat{y}(x) = \frac{\sum_{i \in N_4} k_i y_i}{\sum_{i \in N_4} k_i} = 0.0367$$

where N_4 denotes the nearest neighbor set of size 4.

For $x = 0.4$, we can apply the kernel regression as follows:

$d_1 = 1.3, d_2 = 1.1, d_3 = 0.9, d_4 = 0.7, d_5 = 0.5, d_6 = 0.3, d_7 = 0.1, d_8 = 0.1, d_9 = 0.3, d_{10} = 0.5$

$d_{\max} = 0.5$

$k_1 = -1.6, k_2 = -1.2, k_3 = -0.8, k_4 = -0.4, k_5 = 0, k_6 = 0.4, k_7 = 0.8, k_8 = 0.8, k_9 = 0.4, k_{10} = 0$

$$\hat{y}(x) = \frac{\sum_{i \in N_4} k_i y_i}{\sum_{i \in N_4} k_i} = 0.1967$$

where N_4 denotes the nearest neighbor set of size 4.

$y(0) = 0, y(0.4) = 0.16$

Therefore, the mean square error is $MSE = \frac{1}{2}((0.0367-0)^2 + (0.1967-0.16)^2) = 0.00134$